DCIEM No. CR 2000-108

Decision Making Implications of Multi-Sensor Data Fusion

by

Gary Klein, PhD

Klein Associates Inc.
1750 Commerce Center Blvd. North
Fairborn, OH 45324
E-mail: gary@klein-inc.com

PWGSC Contract No. W7711-9-7592

on behalf of
DEPARTMENT OF NATIONAL DEFENCE

As represented by
Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue West
Toronto, Ontario, Canada
M3M 3B9

DCIEM Scientific Authority:
Name: Dr. Joe Baranski
Telephone no. (416) 635-2146

Date of Preparation: Sept 15, 2000

E-mail: gary@klein-inc.com

DEPARTMENT OF NATIONAL DEFENCE – CANADA

© HER MAJESTY THE QUEEN IN RIGHT OF CANADA (2000)
as represented by the Minister of National Defence

© SA MAJESTE LA REINE EN DROIT DU CANADA (2000)
Défense nationale Canada

102-01023/

Abstract

The use of multi-data sensor fusion to support Operations Room Officers in Halifax Class frigates offers an important potential to improve performance in detecting and identifying tracks. However, there is a risk that the fusion algorithms may also interfere with performance. This report used a Naturalistic Decision Making perspective to identify ways that fusion algorithms might create barriers to the development and application of expertise. One specific concern was that the level of uncertainty might not be reduced, but rather would be shifted from one type of uncertainty to another. Operators using fusion algorithms may have to confront uncertainties about the rationale for system outputs, and may have more difficulty in gaining access to critical data. Another concern was that a system that was efficient during normal operations would become inefficient during non-routine and anomalous events. Recommendations were made for additional research on the relationship between fusion algorithms and expertise. This report takes a deliberately critical position on multi-sensor data fusion (MSDF) because of concerns about its effects on expertise. However, the benefits of MSDF are clear, as is the need to find ways to develop strategies that will make fusion algorithms less disruptive and more valuable in improving performance.

Résumé

L'utilisation de la fusion de données de capteurs multiples à l'appui des officiers du centre des opérations à bord des frégates de la classe Halifax a beaucoup de potentiel pour améliorer les performances de détection et d'identification des cibles. Cependant, les algorithmes de fusion risquent également de nuire aux performances. Le présent rapport a fait appel à une perspective de prise de décision naturaliste pour déterminer les façons dont les algorithmes de fusion risquent de freiner l'élaboration et l'application de l'expertise. Une des préoccupations particulières était que le niveau d'incertitude ne puisse pas être réduit, mais qu'on passe plutôt d'un type d'incertitude à un autre. Les opérateurs utilisant des algorithmes de fusion risquent d'avoir à confronter des incertitudes quant à la raison d'être des sorties du système et ils risquent d'avoir plus de mal à accéder aux données critiques. Une autre préoccupation était qu'un système qui était efficace en conditions de fonctionnement normales devienne inefficace dans des circonstances anormales. On a fait des recommandations en vue de recherches additionnelles sur la relation entre les algorithmes de fusion et l'expertise. Le présent rapport adopte un point de vue délibérément critique à l'égard de la fusion de données de capteurs multiples (MSDF) en raison des préoccupations quant à ses effets sur l'expertise. Toutefois, les avantages de la MSDF sont nets, de même que le besoin de trouver des moyens d'élaborer des stratégies qui rendront les algorithmes de fusion moins disruptifs et plus propices à améliorer les performances.

This page intentionally left blank.

Executive summary

The use of multi-data sensor fusion to support Operations Room Officers in Halifax Class frigates offers an important potential to improve performance in detecting and identifying tracks. However, there is a risk that the fusion algorithms may also interfere with performance, particularly in situations involving missing information, untrustworthy information, and anomalous information. This report used a Naturalistic Decision Making perspective to identify ways that fusion algorithms might create barriers to the development and application of expertise. One specific concern was that the level of uncertainty might not be reduced, but rather would be shifted from one type of uncertainty to another. Operators using fusion algorithms may have to confront uncertainties about the rationale for system outputs, and may have more difficulty in gaining access to critical data. Another concern was that a system that was efficient during normal operations would become inefficient during non-routine and anomalous events. Recommendations were made for additional research on the relationship between fusion algorithms and expertise. This report takes a deliberately critical position on multi-sensor data fusion (MSDF) because of concerns about its effects on expertise. In particular, it is argued that MSDF can reduce performance by interfering with the application of expertise, interfering with the expertise of the team, slowing the rate of learning, and potentially teaching dysfunctional skills. Although, the potential benefits of MSDF are clear and have been well documented, there is a clear need develop strategies that will make fusion algorithms less disruptive and more valuable in improving performance.

G. Klein. 2000. Final Report: Con-fusion Algorithms. CR 2000-108 Defence and Civil Institute of Environmental Medicine.

Sommaire

L'utilisation de la fusion de données de capteurs multiples à l'appui des officiers du centre des opérations à bord des frégates de la classe Halifax a beaucoup de potentiel pour améliorer les performances de détection et d'identification des cibles. Cependant, les algorithmes de fusion risquent également de nuire aux performances, en particulier dans des situations où il y a de l'information manquante, de l'information peu fiable et de l'information anormale. Le présent rapport a fait appel à une perspective de prise de décision naturaliste pour déterminer les façons dont les algorithmes de fusion risquent de freiner l'élaboration et l'application de l'expertise. Une des préoccupations particulières était que le niveau d'incertitude ne puisse pas être réduit, mais qu'on passe plutôt d'un type d'incertitude à un autre. Les opérateurs utilisant des algorithmes de fusion risquent d'avoir à confronter des incertitudes quant à la raison d'être des sorties du système et ils risquent d'avoir plus de mal à accéder aux données critiques. Une autre préoccupation était qu'un système qui était efficace en conditions de fonctionnement normales devienne inefficace dans des circonstances anormales. On a fait des recommandations en vue de recherches additionnelles sur la relation entre les algorithmes de fusion et l'expertise. Le présent rapport adopte un point de vue délibérément critique à l'égard de la fusion de données de capteurs multiples (MSDF) en raison des préoccupations quant à ses effets sur l'expertise. En particulier, il est avancé que la MSDF risque de dégrader les performances en gênant l'application de l'expertise, en nuisant à l'expertise de l'équipe, en ralentissant le taux d'apprentissage et en enseignant éventuellement des aptitudes dysfonctionnelles. Bien que les avantages potentiels de la MSDF soient nets et bien documentés, il existe un besoin évident d'élaborer des stratégies qui rendront les algorithmes de fusion moins disruptifs et plus propices à améliorer les performances.

G. Klein. 2000. Final Report: Con-fusion Algorithms. CR 2000-108, Defence and Civil Institute of Environmental Medicine.

Table of contents

Abstract i

Résumé i

Executive summary iii

Sommaire iv

Table of contents v

List of figures vi

List of tables vi

Acknowledgements vii

Introduction 1

Description of Multi-Sensor Data Fusion (MSDF) 2

Naturalistic Decision Making 5

Framing the Question of How MSDF Affects Uncertainty 10

Under What Conditions Can Data Fusion Increase Uncertainty and Reduce Performance? 14

Recommendations: Making Data Fusion Work 17

Conclusions 20

References 22

List of tables

Table 1. Uncertainty Matrix 12

List of figures

Figure 1. Recognition-Primed Decision (RPD) model 7

Figure 2. Data into understanding 8

Figure 3. Factors in the use of MSDF that could reduce performance 14

Acknowledgements

I would like to thank Joe Baranski, Marvin Thordsen, and Amelia Armstrong for their reviews of earlier drafts, and their suggestions.

Introduction

The purpose of this report is to describe some of the conditions under which multi-sensor data fusion (MSDF) algorithms can add to uncertainty and reduce performance, as well as conditions where data fusion can support skilled task performance. One concern is that the techniques of data fusion may not be entirely consistent with the requirements for skilled judgments on the part of the operator. The solution is not to reject the use of data fusion, but rather to learn the boundary conditions for effective use, and to find strategies for successfully applying data fusion.

The report relies on the Naturalistic Decision Making framework (e.g., Zsambok & Klein, 1997) to examine MSDF. The Naturalistic Decision Making framework studies how people use strategies such as pattern matching to make decisions under time pressure and uncertainty. The framework attempts to understand how people use their expertise to make decisions in field settings. This report takes a critical tone regarding MSDF, attempting to identify potential problems in the way MSDF may interfere with the application of expertise. This should not be interpreted as criticism of the MSDF concept itself. Clearly, there is a great benefit to be obtained from MSDF. The critical tone is intended to encourage a broader evaluation of the effects of MSDF.

The increasing use of information technologies has created a potential dilemma. Previously, commanders were frustrated because of how little information they had. Now, we may be seeing a situation where commanders are frustrated because of how much information they have. The data provided by all the available sensors can easily overwhelm any decision maker. In order to make sense of this information in real time, we need to use computer-based methods to link different data elements whose relationship might escape the notice of the decision maker, and to synthesize higher-level inferences. Data fusion algorithms can accomplish these functions.

But how useful will these algorithms be? The dilemma here is that we may not be reducing the uncertainty facing the commander. Instead, we may just be shifting that uncertainty and re-packaging it. Formerly, the commander experienced uncertainty due to missing information that was not available. Now, the commander may still experience uncertainty if the data have been collected but cannot be located and retrieved when needed. Uncertainty may also arise if the commander has a poor mental model of the way the fusion algorithm works, and is unsure about how to interpret the outputs.

A fusion algorithm is a software program that can identify data elements that are related, and synthesize these to add meaning. For example, if a radar operator has a contact at a certain bearing, and another radar operator has a contact on the same bearing, and the electronic warfare supervisor is getting some emissions at the same bearing, these can be synthesized to identify a single track, and to ascribe certain characteristics to that track. When making use of this capability, the commander and the staff do not have to continually compare notes. The fusion algorithm can accomplish the inferences needed and save the effort of sifting through a great deal of redundant data.

Fusion algorithms therefore offer a valuable means of coping with data overload. And when they work, they can be very beneficial. But that is true of information technologies in general. What happens when they don't work as advertised?

An example may be useful. During the USS VINCENNES shootdown incident in 1988, the VINCENNES identified a track, and assigned it the track number 4474. Shortly thereafter, the USS SIDES also identified a track, and assigned it the track number 4131. (There was an error here in that the SIDES used a track number from the block assigned to the VINCENNES.) The computer system determined that these two tracks were in fact the same aircraft, and gave them a single track number. This is a form of data fusion. Unfortunately, it chose the track number assigned by the SIDES, and the action was taken several minutes into the incident. At a critical point in the incident, the captain of the VINCENNES asked, "What is 4474 doing?" He used the original, now incorrect, track number. Several members of his staff appear to have realized the error, and determined that the track of interest was still climbing. Other members of his staff may have followed his inquiry literally. Inasmuch as the original track number had since been reassigned to another aircraft a few hundred miles away that was in the process of descending, they reported that track 4474 was descending. (This type of assignment has since been changed.) The result was confusion in the Combat Information Center of the VINCENNES, and the decision to engage the track. There were other factors involved; this ambiguity over track numbers was one of many contributors to the shootdown. Nevertheless, the incident illustrates how data fusion can inadvertently interfere with performance.

The objective of this research effort was to draw on knowledge of human decision making in order to identify human factors associated with the paradox that the means of reducing uncertainty through multi-sensor data fusion are themselves a source of uncertainty. We need to examine some of the complexities that must be handled in order to forge an effective working relationship between operators and the computational algorithms. We also need to identify useful strategies for representing the outputs of data fusion technologies so that they can have greater utility to issues involving the use of multi-sensor data fusion (MSDF) technology for the Operations Room Officer (ORO) of the Canadian Halifax Class frigate, which is very roughly analogous to a U.S. AEGIS cruiser. The ORO is analogous to the Tactical Action Officer (TAO) on an AEGIS cruiser. It is important to determine if MSDF creates novel problems for the ORO.

Description of Multi-Sensor Data Fusion (MSDF)

Data fusion. The function of data fusion is to combine different data to generate a higher-order inference. If we consider the goal of MSDF to be the generation of inferences that are unfailingly accurate and reliable, we are being unrealistic. No sensory system is capable of that level of accuracy, not even our own sensory apparatus. Consider some basic processes covered in Psychology 101 courses. At a perceptual level, the sensory apparatus achieves data fusion when a baby learns object constancy. The sensory apparatus achieves data fusion when information from different rod cells in the eye is combined to identify an object. The sensory apparatus achieves data fusion when the firing of different cone cells in the retina, each attuned to different wavelengths of light, is interpreted as an object with

shading. And even under these conditions, errors are made. Objects are misidentified. Size is mistaken for distance. Sensory illusions are easily demonstrated. Thus, we are used to working with a somewhat unreliable mechanism that provides us with fused data and does not permit us to decompose the identification. (We cannot trace a visual identification to the individual retinal cells.) Signal Detection Theory has shown how sensitivity interacts with thresholds, how signals cannot be easily distinguished from noise, and how our experience and expectations dictate the threshold for detecting signals. Therefore, our perceptual systems have not solved the problems of data fusion. We have learned to live with the limitations, and to manage these successfully most of the time.

The same applies at the team level. In the current organizational configuration, without MSDF, the ORO does not receive data directly, but usually receives it from subordinates, the Sensor Weapons Controller (SWC), or the Assistant SWC, who get data directly from the tracker/plotter. Any time information handoffs are required, the possibility for error increases. Thus, an ORO may want to increase the gain, to increase the chance of detecting a target, whereas the tracker/plotter may be skeptical, because increasing the gain will also increase noise, and will reduce the chances for detection. The advantage of using tracker/plotters includes a way to reduce workload for the ORO, a way to increase expertise for the task of target detection, and a way to enable the ORO to ask about the rationale for detection and identification decisions, assuming that the tracker/plotters (and other subordinates) can explain their reasoning.

However, we must recognize that the splitting of functions across different individuals itself increases uncertainty and the potential for errors and confusion. Klein, Armstrong, Woods, Gokulachandra, and Klein (2000) have studied common ground errors in command posts, and describe how easy it is for interpretation errors to emerge and proliferate. Thus, the baseline case itself, without any fusion algorithms, contains types of uncertainty that MSDF could diminish. In the following sections, I will raise some critical concerns about the use of MSDF. However, it should be clear that there is no "ideal" state, and that the pre-MSDF, has its own weaknesses.

Multi-Sensor Data Fusion (MSDF) for detecting and identifying tracks. Smith and Erickson (1991) define MSDF as "the integration and extraction of desired information from data obtained by two or more sensors or by one sensor operating in two or more modes." Elmenreich (2000) explains that MSDF is needed because of drawbacks from employing individual sensors. These include sensor deprivation (breakdown in a critical sensor), limited spatial or temporal coverage, imprecision, and uncertainty (arising from cases where the sensor cannot measure all relevant attributes, or where the observation is ambiguous). Elmenreich sees the potential advantages of data fusion as being: robustness and reliability, extended spatial and temporal coverage, increased confidence, reduced ambiguity and uncertainty, robustness against interference, and improved resolution. Later sections of this report will question whether there are gains in increased confidence and reduced ambiguity.

There are many nuances and methods for MSDF. For the purposes of this report, we will be examining the case where an algorithm uses sensor data (radar, Identification Friend or Foe [IFF], Link-11) to generate estimates of the current and future positions for each relevant object. In addition, at this level the MSDF algorithm

would generate inferences about the identity and key attributes of objects. Basically, the MSDF at this level is using speed, direction, and bearing to generate its inferences.

MSDF distinguishes signals from noise, and objects from clutter. There is a priority placed on identifying all the targets that exist, even though this may generate spurious contacts.

A higher level of analysis might include using the outputs from MSDF to generate a picture of the situation, including friendly and enemy groups and the intents of the enemy groups. Still higher would be an attempt to assign resources and generate plans. In its current form, MSDF is just attempting to indicate possible tracks, and possible threats. It is not trying to infer enemy intent, to interpret what the Rules of Engagement (ROE) permit, and so forth. Some would argue that it is neither realistic nor responsible to have automated agents take on the responsibilities of inferring intent and interpreting ROE, but this report is only directed at the simplest level of fusion, as described above. This report is restricted to the use of Above Water Warfare (AWW), Under Water Warfare (UWW), Link-11, and Wide Area Pictures to provide inputs to a single integrated Maritime Tactical Picture. The AWW sensor suite includes short and long-range radar, IFF, Electronic Support Measures (ESM), and Link-11.

The Data Fusion and Resource Management (DFRM) Group at DREV (Defence Research Establishment Valcartier) has produced a generic MSDF system (Roy & Bossé, 1997). It is recognized that the capability of displaying the outputs of a MSDF system, using advanced display technologies, will require additional effort to conceptualize the Human-Computer Interface (HCI) issues, such as whether to present the tracks generated through MSDF alongside tracks from more conventional sources, or whether to distinguish these tracks and permit further scrutiny and observation. For example, combining chart, sensor, and combat information results in more complex displays, and such complexity can itself interfere with the ability to search for targets. Thus, the designer is faced with the choice of obscuring the details of how inputs are fused (thereby hiding information from the operator) or providing this information at the risk of making the search more difficult, and negating some of the value of the fusion. In addition, as McArthur, Ashley, Bossé, and Beevis (1999) have noted, there may be a need to make the errors of MSDF more obvious to the user, to permit stronger collaboration between the decision makers and the system. This would require that the system alert the human when data are in conflict, reducing confidence in the recommendation of the system.

MSDF addresses individual objects (tracks), either by associating signals with known tracks or by using signals to generate new tracks. The Level-1 processing is aimed at providing an estimate of the current and future position for each (hypothesized) track, and with generating inferences about the identity of a track (object) and some of its key attributes and properties. Radar data (both air and surface), ESM, sonar, Link-11 data, intelligence, and encyclopedic information (such as ferry schedules) are correlated, using mathematical filters. A track is composed of kinematic properties (such as position, course, speed, and acceleration), non-kinematic properties (such as type of target, identity, radar cross-section, IR signature), and the probability of the track. The MSDF attempts to estimate the absolute likelihood that the track exists. MSDF also suppresses clutter in various ways, to eliminate uninteresting objects from the display.

A system such as MSDF will allow the ORO to handle more messages, spot important tracks more quickly, and perhaps achieve the same or higher performance with a smaller cadre

of support staff. MSDF will help the ORO function in the tighter confines of littoral, rather than blue water settings, against high-speed weapons, against a background of large numbers of contacts (typical of littoral settings), with difficult identifications (e.g., weapons systems that can be carried aboard non-conventional surface craft), with the high time pressure typical of littoral settings. The ORO will have to perform his functions in the context of a high degree of communications and a wide range of communications systems, and in the context of geography (e.g., hills) and weather that can degrade sensors and communications. MSDF can be very helpful in environments such as these.

This will only work if the MSDF delivers what is needed, rather than adding to the confusion. By way of analogy, commercial airliners rely on Flight Management Systems (FMS), but some carriers have directed their pilots to turn off the Flight Management Systems in cases of emergency or situations requiring replanning and workarounds. The FMS just adds to the chaos, because as long as it is engaged, the pilot may not be able to understand why it is reacting in certain ways, or what it will do next. This analogy illustrates how fusion algorithms could interfere with performance. One reason pilots may turn off FMS is that during times of stress, workload becomes high and the additional workload needed to reconfigure the system is too much to handle. Intelligent support systems should reduce workload and stress, but may be least helpful during times of stress. Further, without a good mental model of how the FMS works, the pilots run the risk of entering the wrong data, or of generating unintended consequences. These are not issues of trust. They illustrate how automated aids, such as MSDF, may require a rich mental model of the system logic. The requirements for such a deep understanding may not be apparent during routine operations, and may only emerge clearly during workarounds.

Naturalistic Decision Making

The perspective of this report is to draw on the Naturalistic Decision Making (NDM) framework to examine some of the human factors issues affected by the use of MSDF algorithms. This effort is purely at a conceptual level. There was no opportunity to conduct observations or interviews. The content has been developed through the analysis of several reports provided by DCIEM, plus the application of research on the Tactical Action Officer position of the AEGIS cruiser, (Kaempf, Klein, Thordsen, & Wolf, 1996); this work, using Cognitive Task Analysis, was conducted more than 10 years earlier. This report also draws from the discussion of the implications of NDM for information technology presented by Klein (1997), in a project sponsored by the Air Force Research Laboratory.

There are several reasons for drawing on the NDM perspective in this effort.

First, the MSDF environment is marked by time pressure, high stakes, multiple players, dynamic settings, uncertainty, ill-defined goals, and organizational constraints. These are all defining features of NDM (Orasanu & Connolly, 1993) and are all factors that are difficult to impose in tightly controlled laboratory conditions.

Second, the MSDF algorithms are designed to improve situation awareness. In the NDM framework, situation awareness is more important than deciding between courses of

action. This is because pattern recognition skills usually provide people with a sense of how to respond in a situation, once they know what is going on.

Third, NDM is focused on macrocognition, on processes such as problem detection (Klein, Pliske, Crandall, & Woods, 1999b) uncertainty management (Schmitt & Klein, 1996) use of leverage points to build new options (Klein & Wolf, 1998); mental simulation and storytelling (Klein & Crandall, 1995); replanning (Klein, Wiggins, & Schmitt, 1999); and disturbance management (Woods, 1994). These processes occur in a time scale of seconds, minutes or longer. In contrast, microcognition usually addresses phenomena on a time scale of tenths or hundredths of a second (Klein, Klein, & Klein, 2000). MSDF needs to support these macrocognitive processes. Several of them, problem detection and uncertainty management, will be discussed in more detail below.

Fourth, expertise is a key factor for NDM. NDM can be defined as the study of how people use their expertise to make decisions in field settings (Klein, 1998; Zsambok & Klein, 1997). This contrasts with studying laboratory tasks presented to subjects who are novices. The NDM framework contrasts to the use of analytical decision models, which concentrate on ways to calculate probabilities and utility. Expertise is a function of pattern recognition, perceptual skills, a sense of typicality (based on associative learning), and the richness of mental models. This contrasts with views of expertise as the accumulation of rules, of the learning of analytical methods for interpreting data. Klein and Militello (2000) have distinguished between two aspects of expertise: things that experts know, and ways that experts apply this knowledge. Klein and Militello identify several forms of knowledge that experts have: mental models (causal accounts of how things happen in situations), perceptual skills (to make fine discriminations and detect patterns), a sense of typicality (built up from associative learning opportunities), a repertoire of routines (ways of making things happen) and declarative knowledge. Klein and Militello (2000) also discusses how experts can use this knowledge: to spot leverage points, to form expectancies, to project future events, to diagnose previous events, to detect anomalies, to perform workarounds, to make time/distance judgments. There are many other types of applications as well. In thinking about fusion algorithms, it may be helpful to consider the types of expertise they will support, the types of expertise they may replace, and the types of expertise with which they may interfere. We need to see how MSDF can support or interfere with expertise.

Fifth, NDM has provided us with a range of models of how people make decisions under difficult conditions (Lipshitz, 1993). For this report, we will rely on one of the NDM models as a framework for examining MSDF: The Recognition-Primed Decision (RPD) Model (Klein, 1998; Klein, Orasanu, Calderwood, & Zsambok, 1993). Several of the processes found in macrocognition, such as problem detection and uncertainty management, are relevant to the use of MSDF, and these processes will be discussed within the perspective of the RPD model.

The RPD model describes how people can use their experience to size up situations quickly, and to make decisions without having to generate and compare different courses of action. It is relevant to the study of MSDF because we may want fusion algorithms to support this use of expertise, and we may be able to evaluate fusion algorithms by seeing if they interfere with the use of expertise in decision making.

The basic tenets of the RPD model are that people who have gained experience in a field can recognize situations as typical, which includes an appreciation for the plausible goals in the situation, and the courses of action that are likely to be successful. The first course of action considered is usually effective. There are three variations of the model. The simple case is where recognition of a situation occurs and, perhaps due to time pressure, the decision maker simply carries out the first option that is generated. A more complex case is where the decision maker uses mental simulation (Klein & Crandall, 1995) to evaluate the course of action, rather than contrasting it with other courses of action.

For the purposes of this report, we will only consider the version of the RPD model in which the decision maker has to deliberate about the nature of the situation (see Figure 1). The reason for concentrating on this variation is that it matches the function of MSDF, to support situational understanding, rather than to assist an operator in choosing between different courses of action. As shown in Figure 1, in the easy cases, the pattern is clear and the decision maker understands the dynamics of the situation. The RPD model asserts that this situation assessment includes a recognition of plausible goals, relevant cues, expected events, and likely courses of action. In some cases, however, there may be more than one interpretation of the situation, or there may not be any plausible interpretations. In addition, the decision maker may be struck by some anomaly, perhaps a violation of an expectancy, so that the initial interpretation of the situation is brought into doubt. The reactions of the decision maker include trying to gather more information, by attending to the features of the situation to see if an analysis can produce a judgment, or by trying to build a story and using the discrepancies facing the alternative stories to gauge the plausibility of alternative interpretations.

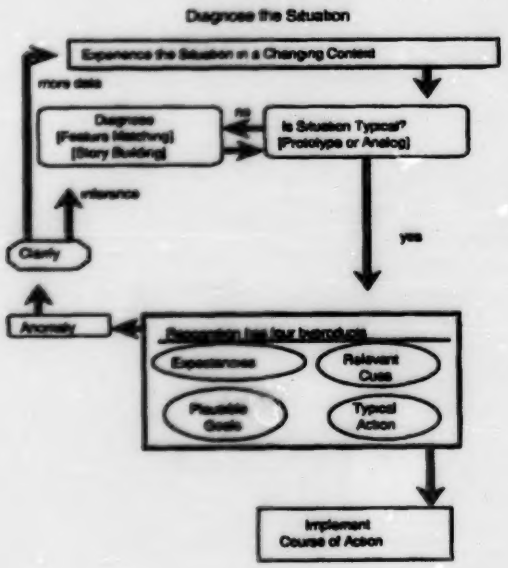

Figure 1. Recognition-Primed Decision (RPD) model

This discussion of how expertise drives naturalistic decision making will provide a basis for the later examination of how MSDF can support or interfere with expertise.

Data-Information-Knowledge-Understanding. As shown in Figure 1, the RPD model asserts that skilled decision makers experience a situation and recognize what is going on. Often, this recognition is accomplished through pattern matching. One temptation is to support decision makers by providing them with an interpretation, so that they do not have to waste time sorting through data. Even without MSDF, command post procedures often try to make this transformation, shown in Figure 2. There are accounts of managing data overload that suggest that low level personnel perform the initial filtering of data, and gradually transform these data into information, and then knowledge, and finally, understanding. Once the inputs have been transformed into understanding, they can be shown to the decision makers.

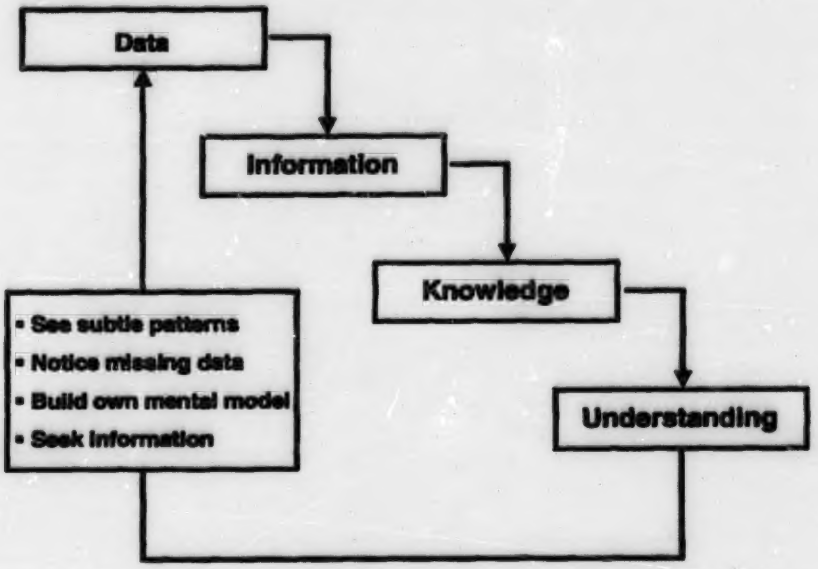

Figure 2. Data into understanding

This view has some severe limitations. Klein (1997) has argued that the transition from data to understanding is not linear. In the field, we find that expertise of the decision makers enables them to see patterns, notice missing data, and make better sense of the data than subordinates. Finally, experts want to build their own mental models, rather than depend on subordinates. Figure 2 describes how experts use their connection to data. The implication is that there is a danger for fusion algorithms to distance skilled decision makers from the data, and thereby run the risk of diminishing their judgment and their confidence.

Problem detection. The RPD model in Figure 1 shows that when an expectancy is violated, and an anomaly is identified, then the process of situation assessment moves into a deliberative mode. The pre-conscious pattern matching is replaced by active attempts to gather more information and to make sense of the inconsistent observations. Therefore, MSDF must support these activities, and not just present interpretations to the decision makers. The system has to make it easy for the operators to spot anomalies.

Problem detection is often critical in natural settings, but it has been rarely studied under laboratory conditions because experiments usually begin by giving subjects problems, rather than letting subjects discover or detect problems. Klein, Pliske, Crandall, and Woods (1999a) have argued that problem detection is usually not a simple accumulation of discrepancies. Instead, it depends on a re-conceptualization of a situation. There are times when anomalous cues may force a person to realize there is a problem. But in other cases, the only way to spot the cues or patterns is to already be aware that there might be a problem. So, which comes first, the early symptoms, or the early suspicions? Klein et al. suggested that in some cases, it might be the symptoms, in other cases, the suspicions, and in other cases, the two may not be different: detecting the cues is the same thing as re-conceptualizing the situation. Further, Klein et al. distinguished between the fault itself, the symptoms of that fault, and the sensor data conveying those symptoms. The system is representing sensor data, but must do so in a manner that permits inferences about the symptoms and the fault. The relevance of this work to fusion algorithms is that operators must be able to spot early signs of threats, and also spot anomalies in the use of the MSDF algorithm.

Uncertainty management. Sometimes, more than one interpretation of a situation is possible. Figure 1 shows several strategies for dealing with this, including attempts to gather more data, attempts to contrast features of the different accounts, and attempts to build and evaluate stories. We can expand on this repertoire.

Schmitt and Klein (1996) have studied the management of uncertainty in military environments. They have distinguished between several different aspects of uncertainty (e.g., uncertainty due to missing data, uncertainty due to lack of confidence in data, uncertainty due to ambiguity and conflicts between data elements). They also identified a range of strategies decision makers use in managing uncertainty. This range of strategies may be useful in thinking about how decision makers will react to the uncertainties reduced by fusion algorithms, or those uncertainties created by fusion algorithms. The strategies for managing uncertainty include the following:

- Gathering more data. This is not always a good strategy, if it sacrifices time.
- Making assumptions to fill data gaps. This is risky, but usually necessary.
- Waiting for the situation to resolve. Again, this sacrifices time.
- Ignoring the uncertainty and pressing on. This is impulsive, but there are times when it is necessary. Data suggest that as people develop wisdom, they learn that uncertainty is inevitable. We also believe there are individual differences in the ability to tolerate uncertainty.

- Plan for the worst plausible case. Here, the uncertainty is managed by making sure the plan is sufficiently cautious to handle setbacks.
- Increasing the robustness of the plan. Similar to the previous point, the strategy is to be more cautious by taking actions such as increasing reserves.
- Being opportunistic. Here, instead of trying to build a detailed plan, one accepts the uncertainty and prepares to improvise as the situation develops, pursuing the lines that are working out the best.
- Shaking the tree. The decision maker resolves the uncertainty by taking an action that requires a reaction, thereby increasing the predictability of the subsequent events.

For a MSDF to be effective, it must support these types of strategies. For example, MSDF should not hinder system operators who encounter anomalies and try to rapidly re-plan. They have to be able to appreciate the reason for the anomalies, by detecting patterns or by detection deviations from expected patterns. If MSDF obscures this type of pattern recognition, it will interfere with the opportunistic strategy of performing improvisations and workarounds.

This brief review of NDM will serve as a basis for evaluating the effect of fusion algorithms on managing uncertainty.

Framing the Question of How MSDF Affects Uncertainty

In this section, we can explore the question of how to assess the impact of a MSDF algorithm. We can see the inadequacy in the question "Will data fusion increase or decrease uncertainty?" Sometimes fusion will reduce uncertainty, and sometimes fusion will increase uncertainty. Fusion can reduce uncertainty when everything is working well and disparate inputs are linked by MSDF to generate an accurate inference. Fusion can increase uncertainty if the operator has more difficulty in accessing data, or in knowing how to interpret the data.

A better question may be "Under what conditions can data fusion increase uncertainty?" But this question is still too simple. We have to deconstruct it: 1) We have to define what we mean by data fusion; 2) We have to define what we mean by uncertainty; 3) We have to identify the conditions affecting the use of data fusion; 4) We have to define the types of operators using data fusion; 5) We have to consider features of the task; and, 6) We have to take into account how the results of the data fusion are represented.

1) We have to define what we mean by data fusion. In this case, we use the MSDF system of combining short-range radar, long range radar, Identify Friend or Foe (IFF) and Link-11 (which provides combat data from other sources).

2) We have to define what we mean by uncertainty. Schmitt and Klein (1996) have distinguished between uncertainty that is due to missing information (either data not received,

or data received but not readily retrieved), data that are received but not trusted, data that are trusted but are incompatible with other trusted data (anomalies), and data that result in such a complex situational interpretation as to confuse the decision maker. Schmitt and Klein also distinguish between data about the present situation, inferences that can be drawn from these data, and projections from the data to future world states.

These factors can be portrayed as a matrix, shown in Table 1. The table highlights the cells that seem most relevant here. The column reflecting uncertainty that is due to excessive complexity does not seem relevant here. That is because MSDF is not attempting to generate a large-scale synthesis of disparate elements, to produce a big picture. It is instead aimed at detecting and identifying single targets that might be embedded in noise and data overload. But the other columns all seem pertinent, as do all of the rows. This means that there are many different ways that MSDF can contribute to the uncertainty of the decision makers, and we need to be sensitive to all of them.

However, we can identify the cells that are most worrisome. I have identified four key concerns:

(a) Missing data about the current situation. The problem here is where the decision maker is unable to decompose the output of MSDF to find out on what it is based.

(b) Missing inferences and projections. Here, the operator may be confused about why MSDF is not providing inputs, and whether it means that there are data about which the decision maker is unaware, or the system is failing (and therefore unreliable), or there are nuances about the algorithms that were not appreciated, or whether there is really not a connection. For example, two reports of tracks may indicate two separate entities in close proximity, rather than a single entity. A commercial airliner may appear to be squawking Model 2 IFF if there is an enemy fighter shadowing it, trying to evade radar detection.

(c) Perceived unreliability of inferences and projections. Here, the decision maker is unable to figure out the nature of the algorithm being used by MSDF, and is therefore unsure about whether to trust it.

(d) Anomalous information. Here, the outputs of MSDF appear credible, but conflict with other data. Thus, in one of the incidents studied by Kaempf, Klein, Thordsen, and Wolf (1996), a report of an Exocet missile was generated, and an AEGIS cruiser prepared to take the necessary defense actions. But the captain quickly became suspicious of the report, because of the absence of corroborating data that he would have expected, such as confirming signals from other ships that should have been in the path of the missile.

Table 1. Uncertainty Matrix

	Missing Information	Untrustworthy Information	Anomalous Information	Complex Information
Current data	(a)			
Inferences about data	(b)	(c)	(d)	
Projections into the future	(b)	(c)	(d)	

3) We have to identify the conditions affecting the use of data fusion. The more opaque the algorithm, the harder it will be for decision makers to understand it. The more deceptive the adversary or the more complex the situation, the greater the chance for real anomalies to emerge, leaving the decision maker to try to figure out if the system is defective or something unusual is occurring. What happens when MSDF malfunctions? How easy will it be for a decision maker to recognize this as a malfunction? This has been recognized by McArthur et al. (1999) as making the system's errors obvious to the user. However, it is unclear how to achieve that. Certainly we cannot depend on a malfunctioning system to make an accurate assessment of its own degradation in order to warn the decision maker of its unreliability.

In general, the more subtle the unexpected event in the environment, the more easily it will be missed or interpreted away, and the presence of MSDF will just make it easier to explain away unusual findings. The operator can shrug off the conflict by saying, "I guess the system is acting up again. Who knows what is going wrong this time?" Thus, we realize that we cannot simply be concerned with the effect of MSDF on experienced uncertainty. We also have to be concerned about cases where uncertainty is low, but problem detection is slowed or degraded. So the question has to be amended to: "Under what conditions can data fusion increase uncertainty and reduce performance?"

4) We have to define the types of operators using data fusion. Are they the current generation, or the future generations? Are they experts or novices? Are they OROs or radar specialists? There is a temptation to focus on current generation, skilled OROs. But this can be too short-sighted. The U.S. Air Weather Service has found that the introduction of information technologies has resulted in forecasting accuracy today that is slightly lower than it was 20 years ago. A major reason for this is because the skill level of the forecasters has been allowed to decline, and because the ease of relying on computer-generated forecasts has enabled forecasters to move into a passive mode wherein they do not try to build their own expertise.

MSDF may simplify performance to the point where officers do not develop the high skill levels they had needed, preferring instead to trust in the system. Klein (1997) has discussed the possibility that information technology can slow down the learning of important skills, in a positive feedback cycle, where the technology leads the decision maker to trust it, and to spend training time elsewhere, realizing that as the essential skills erode, it is more necessary to trust the system's recommendations, and so forth. Mosier (1997) and Smith

(1995) have discussed the problem of an automation bias, in which operators come to rely on the outputs of a computer-generated system, even though their own judgment is more accurate. Instead of trying to formulate their own judgment, they become content to take the advice of the computer. This effect will be seen to some extent in the first generation of users, but it will play out more strongly over time. So the issue is: "For what time frame?" Are we interested in the immediate users, or the second and third generation users?

5) We have to consider features of the task. Is there an adversary? Is the adversary trying to achieve deception? Does the adversary have the tools to achieve deception? Does the task require vigilance for detecting problems and promptly performing workarounds? In the current case, the task is a vigilance and problem detection task, to notice the presence of a threat as quickly as possible. The common downside of automated alarms, such as from a MSDF input, is the rate of false alarms. Here, operators just stop attending to the annoying system. Therefore, research is needed to determine the relationship between sensitivity and false alarms in using MSDF.

In research on AEGIS systems for the TADMUS program, we gathered incidents in which the Tactical Action Officer wanted the radar operator to turn up the sensitivity, in order to spot a potential threat. However, the radar operator believed that this would be counter-productive, and that the increased sensitivity would add more noise than signal, and resisted the directive. With MSDF, the ORO will not even have this level of control. Even if the algorithm automatically adjusts the gain to the best setting, there is the possibility that the skilled ORO will believe he has a better mental model of atmospheric effects and so forth, and will be uncomfortable giving up this control to the system. Thus, the skilled ORO may not trust the MSDF system, whereas in the next generations, OROs may not try to obtain this level of mental model, and will depend too much on the MSDF.

6) We have to take into account how the results of the data fusion are represented. Some displays may make it easier than others to spot anomalies and to interpret the outputs of fusion algorithms. The truism is that the superior displays are ones that package cues and patterns in ways that match the users' mental models and ordinary strategies for performing the task. The challenge is to understand how users are interpreting and synthesizing data elements, and to establish whether some data elements are more easily synthesized than others. Just because it is computationally possible to combine data elements does not guarantee that it will be cognitively meaningful to perceive the elements as a larger unit.

With regard to the issues facing MSDF, we would suggest that the emergent data be distinguishable from combat data. If the users are unable to identify the source, this would seem to add a potential for confusion and error. Over time, changes can be made, once it is determined if users need to distinguish MSDF inputs from other inputs. Rather than trying to figure out if Naval Tactical Data System (NTDS) symbology can or should be used, the safe strategy would be to begin with the distinction (perhaps a slight modification from the NTDS symbology) and to reconsider later.

An example of a simple display concept was one used by Klinger and Militello (1999), in a display for AWACS weapons directors. To alert the weapons directors to threats, a red circle was placed over tracks that were seen as threatening. Two features used to identify threats were that they were flying at a high altitude, above a given level, and that they were flying at a fast rate, above a threshold level. Therefore, tracks that exceeded both of these

parameters were tagged as threats. Other types of information were also of potential value, but the lack of space on a cluttered screen ruled these out. The simple cue, the red circle, was sufficient. The algorithm was straightforward, the cue was known to be computer-generated, and the weapons directors did not report confusion or uncertainty.

Under What Conditions Can Data Fusion Increase Uncertainty and Reduce Performance?

Now that we have honed the question, it is time to try to answer it. Or, rather, it is time to expand on the answers sketched out above.

Figure 3 describes several ways that MSDF could degrade performance:

- MSDF could interfere with the application of expertise by preventing OROs who are highly skilled from using their experience.
- MSDF could interfere with the application of team expertise.
- MSDF could slow the rate of learning in less-skilled OROs.
- MSDF could teach dysfunctional skills to new OROs, and thereby prevent the attainment of expertise.

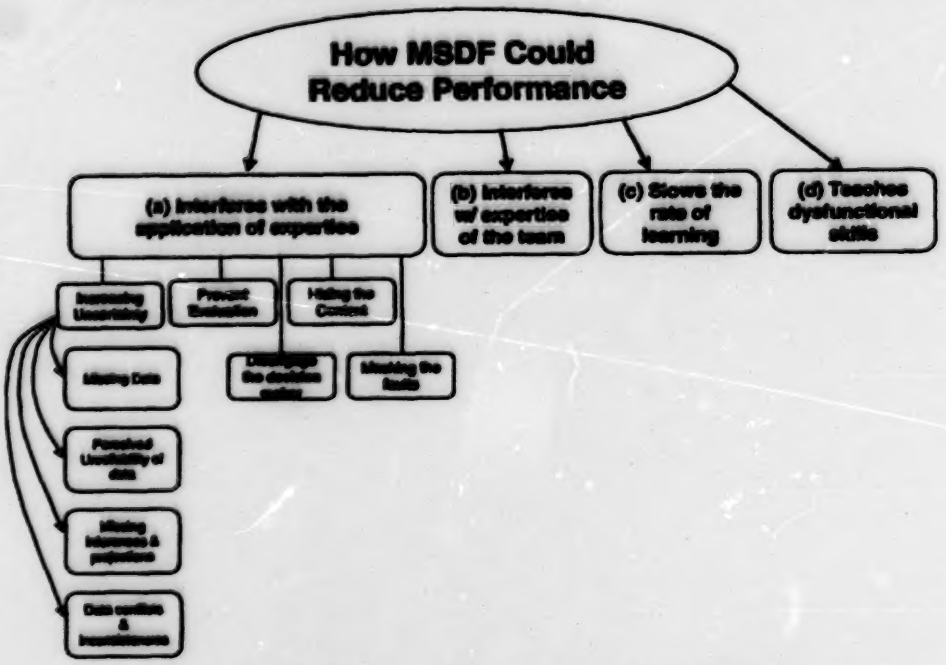
How MSDF Could Reduce Performance
(a) Interferes with the application of expertise
(b) Interferes w/ expertise of the team
(c) Slows the rate of learning
(d) Teaches dysfunctional skills

Figure 3. Factors in the use of MSDF that could reduce performance.

(a) Interfering with the application of expertise. The interference with application of expertise is a problem for most automated systems. Skilled operators can spot anomalies, recognize patterns, and quickly perform workarounds. If a system makes it more difficult to use pattern recognition skills, it can degrade performance in ways that will only be seen during unusual events and breakdowns. Some of the ways that performance disruptions can happen with MSDF are to increase uncertainty, to disengage the decision makers, to prevent evaluation of its outputs, and to hide the context of its outputs.

Using Table 1, we can identify different ways that MSDF could *increase uncertainty*:

- Missing data about the current situation because the ORO cannot decompose the output of MSDF to find out on what it is based.
- Perceived unreliability of inferences because the ORO doesn't understand the MSDF algorithm and cannot trust it.
- Missing inferences and projections because the ORO is confused about why MSDF is not providing inputs.
- Confusion because the outputs of MSDF appear credible, but conflict with other data.

Preventing evaluation—Miller and Militello (1995) described a different problem, involving the requirement to evaluate the output of an algorithm. They studied a case where Artificial Intelligence programs were being used to generate Air Tasking Orders, but in the field the commanders and staff members were not comfortable with the products. The difficulty was that when the Air Tasking Orders were generated manually, they were also being evaluated at the same time. With the automatic generation, there was no opportunity to get a feel for the plan, and the idea of reviewing the massive, detailed Air Tasking Order was impractical. Eliminating the process of plan generation made the staff less aware of the assumptions that had gone into the Air Tasking Order, and the boundary conditions for executing it. Thus, MSDF has the potential to reduce evaluation and scrutiny, providing targets and tracks that are not fully understood.

Automated systems can also *disengage* the decision maker. AWACS weapons directors report that in the "good old days" when they worked on ground control radars, they had to rely on grease pencils to mark the radar returns for each target. This was effortful and tedious, but it also left them truly plugged into the radar picture. In contrast, when they shifted to AWACS and received automatic target generation, their workload went down but they never had the same feel for the status of tracks. On a related theme, MSDF could distract the operators, presenting them with notification of easily detected targets, and creating a false sense of security in which they stop searching actively for difficult to detect targets. Thus, Yeh and Wickens (2000b) report that when they presented automated target cueing, performance improved for detecting the expected targets, but drew attention away from the unexpected targets, which were not cued by the automated program. And, paradoxically, the more reliable the automated cueing, the worse the performance with the unexpected targets. Yeh and Wickens (2000a) reported that sensitivity to detecting new targets decreased when the cueing information was available, compared to conditions under which it was not

available. On the other hand, Dzindolet, Dawe, Pierce, Beck, and Anderson (2000) found a bias towards distrust of automated decision aids. The issue of trust and reliance is not simple (Lee & Moray, 1994). Nevertheless, the possibility exists that operators using MSDF could reduce the effort spent monitoring the situation for unexpected targets that might not be picked up by the fusion algorithm.

A problem can arise when the system *does not show the context* for which the output holds. For example, one decision support system provided outputs about estimated capabilities. However, the display did not show the "reasoning" behind the inference. Thus, a commander may have total air superiority, but if the weather is marked by clouds, rain, and high winds, the air superiority is not going to matter. The commander is not shown whether the algorithm takes weather into account. The skills of the commander, in being able to factor in the effects of context, are discounted. The system needs to present the rationale for the system output, and not just the output itself. Only then can users understand how a change in context will affect the recommendation made by the system. This would also apply to MSDF. If users cannot appreciate the assumptions that go into the track detections and identifications, they will not be able to use their knowledge about weather conditions, jamming, masking, and so forth to judge whether to trust the outputs.

Another way that MSDF could interfere with the application of expertise is *by masking the faults* by making it too easy to discard inconvenient facts. If the MSDF system was not perfect, and if the mental model for the system was weak, operators might find it easy to explain away anomalies as "the system must be acting up again." Thus, performance could decrease because of inappropriate and ill-informed interpretation of the MSDF inputs. Further, if there was ambiguity about the setting of the sensitivity thresholds for the sensors, or factors such as that, it would again be easier to explain away inconvenient facts.

(b) MSDF could interfere with the *expertise of the team*, not just the individual operator. Hammes (2000) describes a U.S. Marine Corps exercise using an automated system to support artillery planning. In the manual system, all missions came in via voice circuits, contributing to an already noisy environment. The artillery liaison officer, upon hearing of an artillery mission, recorded the pertinent data on a 3x5 card that he displayed for the others. The automated system presented the information silently to the artillery liaison officer. But none of the others (e.g., the air officer, naval gunfire officer, and so forth) ever heard of the artillery mission (no audio cues) or saw the 3x5 card (no visual cue), so the teamwork was disrupted rather than improved. With MSDF, the possibility may exist for a similar interference with team coordination.

(c) MSDF could reduce performance by slowing the rate of learning. The problem is that MSDF could slow the rate of learning. Here, the concern is with the next generation of OROs, who have not had experience with finer levels of data, and are more reliant on MSDF. The less skill they have, the more they are likely to rely on MSDF, and the more they rely on MSDF, the fewer skills they develop. This is a vicious cycle. The end point is for the OROs of the future to become heavily reliant on MSDF, and to be unable to perform workarounds in the event of an MSDF malfunction or error.

McDermott, Klein, Thordsen, and Paley (2000) describe an automated support system for the Airborne Laser program, in which the algorithms were calculating whether a target could be hit, and whether there was a risk of damaging friendly assets. The algorithmic

support was needed because of the extreme time pressure. However, in some cases, the system refused to fire at what appeared to be an obvious target. Further, the system did not provide any clues about why it wasn't firing. The operators would rapidly switch to a manual mode. However, they had difficulty in determining if there was a potential for fratricide should the Airborne Laser miss its target, because they had learned to rely on the automated system for alerts about firing conflicts, and had not developed their own, manual strategies for picking up such conflicts. Thus, their experience with the system made it more difficult for them to operate in a manual mode. This is a danger that needs to be avoided for MSDF implementation.

(d) MSDF could reduce performance by teaching dysfunctional skills. Some systems have the potential to develop operator skills that can be counter-productive. For example, performance measurement systems that provide immediate feedback during training can reduce the transfer to the operational environment, because decision makers have not developed skills of obtaining their own feedback (Schmidt & Wulf, 1997). In the previous example involving Airborne Laser, it could be argued that there was some teaching of dysfunctional skills. The operators were learning to manage their attention in the context of an automated task, and were looking in the wrong places when it came time for manual performance. In general, the dysfunctional skills probably involve attention mismanagement. For MSDF, one risk is that operators will develop habits for seeking information that may interfere with performance in a manual mode.

Recommendations: Making Data Fusion Work

This section covers three topics about improving the introduction of MSDF. First, there is a series of recommendations on how to improve the design of systems such as MSDF. Second, there is some discussion of how to evaluate data fusion systems. Third, several research questions are presented.

Design suggestions:

- Provide the ORO with selective access to details. Enable the ORO to select the types of details needed. Thus, instead of a generic "declutter" key, the ORO should be able to declutter in a specific geographical area, or a specific portion of the spectrum. Generic declutters often just produce frustration. The ORO may have a specific idea of what he is looking for, and where that might be found.

- Provide explanations for the information, not just confidence ratings. Instead of having MSDF offer its own judgment of its own ability (a tenuous judgment, because it depends on the fallible system to judge its own fallibility), it might be better to provide the ORO with a means of determining the cues being used, and their weights. (In short, provide the operator with a "rationale" window that pops up and is easily turned off once the operator has taken it into account.) The ORO can then make his judgment of whether to trust MSDF.

- Display the gain control. Instead of just letting the ORO adjust the sensitivity of the sensors, it might be useful to have an adjunct screen displaying the signal-to-noise tradeoff function.

- Support active construction of situation awareness. Instead of teaching the ORO to use and trust MSDF, it might be better to teach the ORO to form his own judgments, and not to trust MSDF. Thus, in training, it might be useful to have a mode or exercise where MSDF generates spurious tracks, so the ORO has to learn how to spot these. It may be a mistake to worry about teaching trust in any system. We do not want the ORO to take a passive role. We want him to be an active partner, using a tool. As an example, a U.S. Navy Electronic Warfare technician, skilled at operating the SLQ-32, once explained that, "the SLQ-32 is a liar." By that, he meant that it sometimes generated spurious contacts. This did not trouble him. He had a good sense of when this might occur, or when the contacts were not credible, and he had an easy routine for clearing the system to see if the contact was again generated (which showed that it wasn't a transient). And that enabled him to use the SLQ-32. He did not trust the system to be infallible. Permit the ORO to control the spotlight of MSDF, perhaps altering the false alarm rate functionally. In some areas, any hint of a track must be detected. In others, early detection is less important. We want to give the ORO the capability of using MSDF for selective attention.

Evaluation strategies:

- In testing MSDF, it might be useful to run scenarios where there are inconsistent cues or tracks, to see how easily the ORO can manage these. A variant here is to run scenarios where the ORO is surprised, as the situation is inconsistent with expectancies. It could be useful to see if the ORO falls into the trap of using MSDF to preserve expectancies by explaining away inconvenient data.

- Another variant would be to impose sensor failures, especially intermittent ones, including transient errors. Researchers could vary the false alarm rate, or impose transmission failures of different sorts. Include scenarios that have a great deal of ambiguity, permitting multiple possible interpretations and trends. Will the ORO just seize on the interpretation offered by MSDF, or will he see alternative explanations for the data?

- When the ORO has access to the same data as MSDF, does the ORO trust his own judgment more or less than MSDF? If he trusts MSDF more, that may be a bad sign.

- Does the ORO have the same confidence in all aspects of MSDF? Thus, if the ORO believes the inferences generated by MSDF about enemy

capabilities, as much as the straightforward generation of tracks, that might signal an uncritical attitude that is dangerous.

- Does the ORO have a good sense of the boundary conditions for MSDF? Is there a reasonable mental model of how MSDF generates its outputs? Researchers could test the ORO's mental model, and evaluate its accuracy.

- Do new OROs show any reduction in expertise about individual cues and their meaning? If so, that may be an unfortunate sign that MSDF diminishes expertise.

- Do OROs use MSDF differently in working against an enemy that is more deceptive, and has more capability for deception? If not, if the use of MSDF is fairly stereotyped, that is a sign that the ORO is moving into a passive mode of just operating the system.

Research strategies. In preparing to field MSDF, it may be useful to run some studies. The evaluation strategies listed in this sub-section can be transformed into controlled experiments. In addition, there are other possibilities for research:

- Studying the existence of an automation bias for situation assessment. Mosier (1997) and Smith (1995) have demonstrated the existence of an automation bias whereby people perform at a lower level if the system generates a recommendation that merely needs to be reviewed, versus having the operator generate a course of action first, and then seeing the system's recommendation. Can we find an automation bias with regard to detecting tracks using MSDF? The study might involve having some trials where operators get the notification from the system, and can also perform their own subsequent examination if they wish, versus a condition where the operators do their scanning first, and only then see if the system has come up with anything. The prediction is that the former condition would disengage operators, reduce vigilance, and result in diminished performance, as was reported by Yeh and Wickens (2000a, b).

- Studying the effect on skill development of using MSDF. The study would contrast operators using MSDF from the beginning of their training with operators receiving training on manual detection/identification. Compared to current operators, operators trained primarily with MSDF should be less capable when MSDF is not available. Furthermore, the operators trained with MSDF should have more difficulty managing anomalies stemming from sensor unreliability. They should be less able to distinguish anomalies based on atmospheric conditions, sensor failures and system faults. The ability to detect and manage anomalies in general should be lower. The MSDF-trained operators should have more shallow mental models, and should be less capable of responding to false alarms.

- Studying the effectiveness of a training program to support MSDF usage. It should be possible to develop a program that exposes ORO's to system failures and unreliabilities, and helps them build a mental model of MSDF. This would provide dexterity in performance workarounds with the system.

- Studying the effect of stress on the use of MSDF. Would stress increase or decrease the acceptance of tracks? Stress can either lead to greater reliance on an automated system or it can lead a person to abandon an automated system (e.g., FMS for commercial aviation). Presumably, this will vary as a function of workload, degree of stress, and expertise of the ORO. Thus, there should not be a simple relationship (i.e., "stress will increase the reliance on MSDF").

- Will operators accept the tracks established by MSDF? This involves factors such as common ground, trust, stress, and mental models. And there should be no simple answer to this question either. In reviewing the literature, we found cases where operators relied too heavily on an automated aid (e.g., Yeh & Wickens, 2000a, b), and cases where they did not trust it sufficiently (e.g., Dzindolet et al., 2000). (Also, see Lee and Moray, 1994.) But, do we want the OROs to accept the tracks generated by MSDF? Wouldn't we want some degree of skepticism? An ORO would maintain some skepticism about track nominations produced by a plotter/tracker, even a skilled one. Klein (1998) described an incident during Desert Storm where a naval officer on the HMS Gloucester detected an Iraqi silkworm launch without any clear cues or characteristics. He was not even able to articulate how he made the judgment, and it appeared to be based on the absence of a signal at a critical point. MSDF would not have substituted for this type of judgment. The question is whether MSDF would have interfered with this type of exercise of expertise.

Conclusions

This report has examined some of the potential difficulties in using a fusion algorithm. No attempt has been made to determine if the overall result of fusion algorithms for detecting and identifying targets will be positive or negative. Rather, the emphasis has been on using a Naturalistic Decision Making perspective to trace the possible difficulties that might arise from a system such as MSDF. The primary concern is that MSDF may be effective in reducing uncertainty and workload when everything works well, but that MSDF may interfere with the development and application of expertise for handling anomalies and workarounds, when an operator may need to look at the data elements more carefully and flexibly.

What is the impact of MSDF on situation awareness, and on the perception of a threat? This relates to whether MSDF is a means of reducing the OROs workload, reducing

the staffing by eliminating some positions, reducing the required skill level of OROs, or a means of extending the OROs capability. If MSDF is easily used and queried, and its boundary conditions are understood, the ORO may be able to use it as a tool, and potentially become more effective in working with MSDF than even working with a tracker/plotter. If MSDF supports the development of skill in an ORO, and the maintenance of an active engagement, then the ORO/MSDF partnership could result in better performance.

References

Dzindolet, M. T., Dawe, L. A., Pierce, L. G., Beck, H. P., & Anderson, B. W. (2000). Misuse of an automated decision making system. Proceedings of the 5th International Conference on Human Interaction with Complex Systems, 81-84.

Elmenreich, W. (2000). An introduction to data fusion (Technical 11/2000). Vienna: Institut f. Technische Informatik Technical University.

Hammes, T. X. (2000). Do no harm. Marine Corps Gazette, 76-80.

Kaempf, G. L., Klein, G., Thordsen, M. L., & Wolf, S. (1996). Decision making in complex command-and-control environments. Human Factors Special Issue, 38, 220-231.

Klein, D. E., Klein, H. A., & Klein, G. (2000). Macrocognition: Linking cognitive psychology and cognitive ergonomics. Proceedings of the 5th International Conference on Human Interactions with Complex Systems, 173-177.

Klein, G. (1997). Implications of the naturalistic decision making framework for information dominance (Contract F41624-94-D-6000 for Armstrong Laboratory, WPAFB, OH). Fairborn, OH: Klein Associates Inc.

Klein, G. (1998). Sources of power: How people make decisions. Cambridge, MA: MIT Press.

Klein, G., Armstrong, A., Woods, D. D., Gokulachandra, M., & Klein, H. (2000). "Being on the same wavelength": The role of common ground in distributed replanning (Technical Report submitted to the Air Force Research Laboratory/HECA under Contract F41624-99-C-6051, Wright Patterson AFB, OH). Fairborn, OH: Klein Associates.

Klein, G., & Militello, L. (2000). The knowledge audit as a method for cognitive task analysis. Proceedings of the 5th Conference on Naturalistic Decision Making: How Professionals Make Decisions (CD-Rom). Stockholm, Sweden: National Defence College.

Klein, G., Pliske, R. M., Crandall, B., & Woods, D. (1999a). Features of problem detection. Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, 1, 133-137.

Klein, G., Pliske, R. M., Crandall, B., & Woods, D. (1999b). Problem Detection. Manuscript submitted for publication.

Klein, G., Wiggins, S. L., & Schmitt, J. (1999). Cognitive aspects of replanning at Army division-level command posts (Final Report Contract DAAH01-99-C-R060 for U.S. Army Aviation and Missile Command). Fairborn, OH: Klein Associates Inc.

Klein, G., & Wolf, S. (1998). The role of leverage points in option generation. IEEE Transactions on Systems, Man and Cybernetics: Applications and Reviews, 28(1), 157-160.

Klein, G. A., & Crandall, B. W. (1995). The role of mental simulation in naturalistic decision making. In P. Hancock, J. Flach, J. Caird, & K. Vicente (Eds.), Local applications of the ecological approach to human-machine systems (Vol. 2, pp. 324-358). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). Decision making in action: Models and methods. Norwood, NJ: Ablex.

Klinger, D. W., & Militello, L. G. (1999). Designing for performance: A cognitive systems engineering and cognitive task analysis approach to the modification of the AWACS weapons director interface: Manuscript in preparation.

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. International Jouranl of Human-Computer Studies, 40, 153-184.

Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 103-137). Norwood, NJ: Ablex.

McArthur, B., Ashley, A., Bossé, É., & Beevis, D. (1999). The command decision aid technology (COMDAT) Project. Paper presented at the TTCP Symposium on Battle Management, San Diego, CA.

McDermott, P., Klein, G., Thordsen, M., & Paley, M. (2000). Representing the cognitive demands of new systems: A decision-centered design approach (Final Report under contract F41624-99-C-6029 for the AFRL/HECI, Wright Patterson AFB, OH). Fairborn, OH: Klein Associates Inc.

Miller, T. E., & Militello, L. G. (1995). Increasing the robustness of AI generated plans (Technical Report No: RL-TR-95-30). Griffiss Air Force Base, NY: Rome Laboratory.

Mosier, K. L. (1997). Myths of expert decision making and automated decision aids. In C. E. Zsambok & G. Klein (Eds.), Naturalistic decision making (pp. 319-330). Mahwah, NJ: Lawrence Erlbaum Associates.

Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 3-20). Norwood, NJ: Ablex.

Roy, J., & Bossé, É. (1997). A generic multi-source data fusion system (DREV Report No. 9719).

Schmitt, J. F., & Klein, G. (1996). Fighting in the fog: Dealing with battlefield uncertainty. Marine Corps Gazette, 80, 62-69.

Schmidt, R. A., & Wulf, G. (1997). Continuous concurrent feedback degrades skill learning: Implications for training and simulation. Human Factors, 39(4), 509-525.

Smith, C. R., & Erickson, G. J. (1991). Multisensor data fusion: Concepts and principles, IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 1, 235-237).

Smith, P. J. (1995). Implications for strategic planning and cooperative problem-solving. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, 1, 102.

Woods, D. D. (1994). Cognitive demands and activities in dynamic fault management: Abduction and disturbance management. In N. Stanton (Ed.), Human factors of alarm design. London, England: Taylor & Francis.

Yeh, M., & Wickens, C. D. (2000a). Effects of cue reliability, realism, and interactivity on biases of attention and trust and augmented reality. Proceedings of IEA 14th Triennial Congress and HFES 44th Annual Meeting, 1, 511-514). Santa Monica, CA: Human Factors and Ergonomics Society.

Yeh, M., & Wickens, C. D. (2000b). Examination of explicit and implicit display signalling on attention allocation and trust calibration. Proceedings of the 5th International Conference on Human Interaction with Complex Systems, 45-49.

Zsambok, C. E., & Klein, G. (Eds.). (1997). Naturalistic decision making. Mahwah, NJ: Lawrence Erlbaum Associates.

DOCUMENT CONTROL DATA SHEET

1a PERFORMING AGENCY Gary Klein, Ph D , Klein Associates Inc , 1750 Commerce Center Blvd North, Fairborn, OH 45324	2 SECURITY CLASSIFICATION UNCLASSIFIED
1b PUBLISHING AGENCY DCIEM	
3 TITLE (U) Decision Making Implications of Multi-Sensor Data Fusion	
4 AUTHORS Klein, Gary	
5 DATE OF PUBLICATION September 15 , 2000	6 NO OF PAGES 34
7 DESCRIPTIVE NOTES	
8 SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY Sponsoring Agency Monitoring Agency Contracting Agency DCIEM Tasking Agency	

9 ORIGINATORS DOCUMENT NUMBER Contract Report 2000-108	10 CONTRACT GRANT AND/OR PROJECT NO W7711-9-7582	11 OTHER DOCUMENT NOS

12 DOCUMENT RELEASABILITY Unlimited distribution
13 DOCUMENT ANNOUNCEMENT Unlimited

14 ABSTRACT

(U) The use of multi-data sensor fusion to support Operations Room Officers in Halifax Class frigates offers an important potential to improve performance in detecting and identifying tracks. However, there is a risk that the fusion algorithms may also interfere with performance. This report used a Naturalistic Decision Making perspective to identify ways that fusion algorithms might create barriers to the development and application of expertise. One specific concern was that the level of uncertainty might not be reduced, but rather would be shifted from one type of uncertainty to another. Operators using fusion algorithms may have to confront uncertainties about the rationale for system outputs, and may have more difficulty in gaining access to critical data. Another concern was that a system that was efficient during normal operations would become inefficient during non-routine and anomalous events. Recommendations were made for additional research on the relationship between fusion algorithms and expertise. This report takes a deliberately critical position on multi-sensor data fusion (MSDF) because of concerns about its effects on expertise. However, the benefits of MSDF are clear, as is the need to find ways to develop strategies that will make fusion algorithms less disruptive and more valuable in improving performance.

L'utilisation de la fusion de données de capteurs multiples à l'appui des officiers du centre des opérations à bord des frégates de la classe Halifax a beaucoup de

potentiel pour améliorer les performances de détection et d'identification des cibles Cependant, les algorithmes de fusion risquent également de nuire aux performances Le présent rapport a fait appel à une perspective de prise de décision naturaliste pour déterminer les façons dont les algorithmes de fusion risquent de freiner l'élaboration et l'application de l'expertise Une des préoccupations particulières était que le niveau d'incertitude ne puisse pas être réduit, mais qu'on passe plutôt d'un type d'incertitude à un autre Les opérateurs utilisant des algorithmes de fusion risquent d'avoir à confronter des incertitudes quant à la raison d'être des sorties du système et ils risquent d'avoir plus de mal à accéder aux données critiques Une autre préoccupation était qu'un système qui était efficace en conditions de fonctionnement normales devienne inefficace dans des circonstances anormales On a fait des recommandations en vue de recherches additionnelles sur la relation entre les algorithmes de fusion et l'expertise Le présent rapport adopte un point de vue délibérément critique à l'égard de la fusion de données de capteurs multiples (MSDF) en raison des préoccupations quant à ses effets sur l'expertise Toutefois, les avantages de la MSDF sont nets, de même que le besoin de trouver des moyens d'élaborer des stratégies qui rendront les algorithmes de fusion moins disruptifs et plus propices à améliorer les performances

15 KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) Multi-sensor data fusion, Naturalistic Decision Making, Individual and Team Performance, Expertise, Decision Support Systems

#514823